5-hydroxymethylcytosine profiling as an indicator of cellular state

5-hydroxymethylcytosine profiling as an indicator of cellular state

DNA methylation is widely studied in the context of cancer. However, the rediscovery of 5-hydroxymethylation of DNA adds a new layer of complexity to understanding the epigenetic basis of development and disease, including carcinogenesis. There have been significant advances in techniques for the detection of 5-hydroxymethylcytosine and, with this, greater insight into the distribution, regulation and function of this mark, which are reviewed here. Better understanding of the associated pathways involved in regulation of, and by, 5-hydroxymethylcytosine may give promise to new therapeutic targets. We discuss evidence to support the view of 5-hydroxymethylcytosine as a unique and dynamic mark of cellular state. These 5-hydroxymethylcytosine profiles may offer optimism for the development of diagnostic, prognostic and predictive biomarkers.

KEYWORDS: cancer ■ cellular state ■ epigenetics ■ hydroxymethylcytosine

Alexander Laird[1,2], John P Thomson[1], David J Harrison[2,3] & Richard R Meehan*[1]

[1]MRC Human Genetics Unit, Institute of Genetics & Molecular Medicine, University of Edinburgh, Western General Hospital,Edinburgh, EH4 2XU, UK
[2]Edinburgh Urological Cancer Group, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, UK
[3]School of Medicine, Medical & Biological Sciences Building, University of St Andrews, St Andrews, KY16 9TF, UK
**Author for correspondence: Tel.: +44 131 332 2471 richard.meehan@igmm.ed.ac.uk*

Epigenetic processes are associated with the regulation of gene expression through the chemical tagging of chromatin and DNA, which can alter their interaction with transcription factors [1]. Fundamentally, the field of epigenetics is concerned with inherited (but potentially reversible) changes in gene expression or cellular phenotype that do not rely on alterations to the underlying DNA sequence [1]. This definition has been broadened to include nonheritable histone modifications and transient changes associated with DNA repair or cell-cycle phases, such that modification of DNA, RNA and histones is broadly viewed as 'epigenetic' in character [2]. From this perspective, dynamic changes in DNA and histone modifications can be observed as embryogenesis proceeds in somatic, stem and primordial germ cells, which correlate with dynamic changes in gene expression [3–5]. These epigenetic signatures are characteristic of a particular cell type as they help define the associated transcriptome; this is most obvious in disease states such as cancer, which exhibit altered epigenetic and transcriptomic profiles [6–10].

DNA modifications are found to exist at cytosine bases in CpG dinucleotide sequences. Historically, the most widely studied modification of DNA is the incorporation of a methyl group onto the fifth carbon of a cytosine base to form 5-methylcytosine (5mC) [11]. This heritable epigenetic mark has been demonstrated to be present in all vertebrates and is proposed to be essential for normal development and cellular survival. The necessity of 5mC is highlighted with the dramatic global loss of 5mC, altered gene expression patterns and embryonic lethality with targeted inactivation of the enzymes responsible for the initiation and maintenance of normal methylation patterns (the DNA methyltransferases [*Dnmt1–3*]) in mice [12–14]. Functionally, promoter 5mC contributes to the regulation of transcription at associated genes and has also been implicated in genomic imprinting, X-chromosome inactivation, tissue-specific gene regulation and retrotransposon silencing [15]. At the same time, potential roles for DNA methyltrasferase enzymes independent of their catalytic activity should not be ignored [16,17].

As originally proposed, DNA methylation was generally viewed as a stable epigenetic modification [18]. The preference of DNMT1 for hemimethylated substrates made it ideally suited to establish stable differentiated states in the absence of genetic mutation [19]. On the other hand, large-scale loss and reaccumulation of 5mC during murine embryogenesis and cellular differentiation indicated that DNA methylation profiles are 'reprogrammable' [4]. The identification of the molecular pathways underpinning DNA demethylation has a checkered history [20]. The recent identification of TET proteins as potential regulators of DNA methylation patterns through the oxidative conversion of 5mC to 5-hydroxymethylcytosine (5hmC) has sparked intense activity in this research area [21]. There has been a growing body of inquiry into the function, dysregulation and potential significance of 5hmC in physiology

Future Medicine part of fsg

and pathology [22,23]. This article will review the technologies for 5hmC detection and the insights this has given us into 5hmC function. This will provide the foundations to discuss tissue-specific profiles and disease-specific changes, which may imply that 5hmC is important in the pathogenesis of cancer or potentially useful as an indicator of disease state.

Discovery & technology

5hmC was described in 1952, having been identified in DNA from bacteriophages and viral DNA [24]. It took 20 years before the modification was first described in mammals, with the isolation of 5hmC carried out in rat, mouse and frog brains, as well as rat livers, at a yield of approximately 20 and 6% of cytosine, respectively [25]. However, these levels appeared too high and were not reproducible and, as such, were not widely accepted [26]. It was after the investigative discovery of 5hmC in murine neuronal cells in 2009 that the interest in this modified base was reignited [27]. While aiming to compare quantitative differences in 5mC between Purkinje and granule cell nuclei, Kriaucionis *et al.* discovered an unusual DNA nucleotide using thin layer chromatography [27]. They were then able to confirm this to be 5hmC with HPLC and mass spectrometry [27]. These techniques allowed accurate quantification of 5hmC, but it was realized that standard techniques to assess 5mC could not differentiate the two [28–30].

Digestion of DNA before PCR with methylcytosine-sensitive HpaII and methyl-insensitive isoschizomer MspI is able to discriminate unmethylated cytosine and 5mC. Differential quantification can also be achieved through PCR or sequencing after bisulfite conversion of cytosine to uracil through deamination, which is inhibited by methylation [31–33]. However, importantly, 5hmC as well as 5mC was shown to completely inhibit HpaII, as well as other methyl-sensitive restriction enzymes and 5hmC is also resistant to bisulfite deamination [29], thus preventing differentiation of 5mC from 5hmC. During bisulfite conversion, 5hmC is converted to 5-methylenesulfonate, which was shown to be less efficiently amplified during PCR, thus raising the possibility of underestimation of these regions [28,34]. These findings highlighted the need for reassessment of previous methylation data to identify the contribution of 5hmC and the need for novel detection methods to allow accurate 5hmC localization, quantification and assessment of function, particularly as 5mC and 5hmC may occur in the same DNA fragments.

An overview of 5hmC detection methods is shown in Figure 1. 5mC enrichment through antibody immunoprecipitation has previously been established [35,36]. Initially, 5hmC polyclonal and monoclonal antibodies were used for semiquantitative purposes, as well as for immunoprecipitation and downstream analysis with PCR, array profiling or sequencing [37]. Bisulfite conversion of 5hmC to 5-methylenesulfonate and antibody immunoprecipitation of the converted base has also been demonstrated [34,38]. In addition to antibody-mediated methods of 5hmC enrichment, several groups have since developed alternative enrichment strategies. Several of these are centered on the selective conversion of the modification into a glucosylated form following incubation with the T4 β-glucosyltransferase and modified dUTP. Following conversion of the 5hmC marks, these sugar-coated motifs can be purified by a variety of techniques such as selective biotinylation and streptavidin pull-down [34,39], or anaylzed further through selective methyl sensitive restriction enzyme digests [40,41]. Despite this, none of these methods allowed accurate localization of 5hmC, which led to the development of single-base quantification with oxidative bisulfite sequencing (oxBS-seq) and TET-assisted bisulfite sequencing (TAB-seq).

oxBS-seq relies on the conversion of 5hmC to 5-formylcytosine (5fC) with potassium perruthenate, which does not modify 5mC or cytosine. This 5fC is sensitive to bisulfite conversion to uracil. As such, standard bisulfite conversion can be performed on untreated DNA, which converts only cytosine to uracil, as well as the potassium perruthenate-treated DNA, which converts both cytosine and 5fC. The two libraries can be subtracted for accurate localization and quantification of cytosine, 5mC and 5hmC [42]. TAB-seq is performed first by glucosylation of 5hmC to prevent oxidation. 5mC is subsequently oxidized with excess TET enzymes to 5-carboxylcytosine (5caC), without conversion of 5hmC. The sample then undergoes standard bisulfite conversion, which converts the cytosine and 5caC (formerly 5mC) to uracil with protection of the modified 5hmC. The TAB-seq library is then compared with standard bisulfite sequence libraries for base-level localization and quantification [43]. However, despite these advances in single-base resolution profiling of 5hmC, these techniques are largely still in development and have not been fully embraced by the epigenomic community. This is likely due to a number of factors: the development

of reliable commercial kits for oxBS-seq and TAB-seq is still ongoing, particularly as TAB-seq requires highly active TET enzymes; second, many interested groups are still limited by the cost of sequencing; oxBS-seq may also be limited by the damage and degradation of genomic DNA by chemical oxidation conditions and repeated bisulfite treatments required for deamination. Nonetheless, these technological advances underpin our greater awareness of 5hmC distribution and fuel intellectual curiosity regarding its potential function and interpretation.

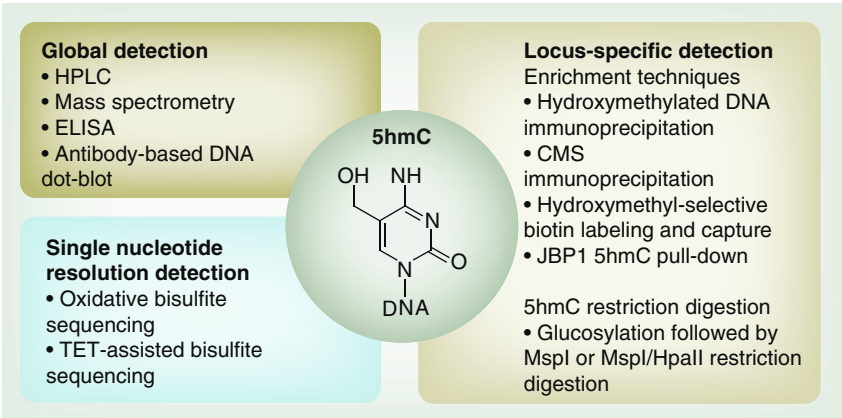

Figure 1. Global, locus-specific and single nucleotide resolution of 5-hydroxymethylation detection methods. HPLC, mass spectrometry , ELISA and DNA dot-blot techniques provide global assessment of 5hmC levels. Locus-specific assessment of 5hmC can be achieved through antibody immunoprecipitation of 5hmC directly or CMS following bisulphite conversion of 5hmC to CMS. Other locus-specific detection methods rely on the glucosylation of 5hmC with subsequent pull-down techniques, selective labeling with biotin and chemical capture or selective methyl-sensitive restriction enzyme digestion approaches. Single nucleotide resolution profiling of 5hmC can be achieved through two methods, either oxidative bisulfite sequencing or TET-assisted bisulfite sequencing. oxBS-seq relies on the oxidative conversion of 5hmC to 5-formylcytosine by potassium perruthenate, which is susceptible to bisulfite conversion such as unmodified cytosine. While during TAB-seq, 5hmC is protected from bisulfite conversion and 5mC is made sensitive through glucosylation and oxidation to 5hmC and 5-carboxylcytosine, respectively. These unique oxBS-seq and TAB-seq libraries can be subtracted from standard bisulfite-sequence libraries to give profiling of relative 5hmC and 5mC.
5hmC: 5-hydroxymethylation; CMS: 5-methylenesulfonate.

Function & distribution

In 2009 Tahiliani and colleagues identified the TET enzymes and their role in the methylation cycle as potential demethylases, based on their similarity between mammalian 5mC and base J in kinetoplasids (a group of unicellular protozoa), which is involved in gene silencing [21]. They used a bioinformatics approach to identify mammalian homologs of the proteins involved in these conversions (JBP1 and JBP2: 2-oxoglutarate [2OG]- and Fe[II]-dependent oxygenases), hypothesizing that if present, these would be involved in 5mC modification [44]. This led to the discovery of the paralogous human proteins TET1, TET2 and TET3 [21], which were also demonstrated to be 2OG- and Fe(II)-dependent enzymes. Overexpression of TET enzymes in human and mouse cells resulted in increased 5hmC and decreased 5mC detection, with a depletion of TET enzymes associated with a reduction of 5hmC [21,45,46]. These findings suggested that the TET enzymes may cause oxidation of 5mC to 5hmC and led to studies investigating the function of 5hmC. A number of groups hypothesized that 5hmC was an intermediate in a DNA demethylation pathway [47,48].

The mechanisms of 5hmC involvement in demethylation were proposed to be both active and passive. In the passive model, the usual mechanism for maintenance of 5mC is disrupted. During cell division, DMNT1 is able to recognize hemimethylated DNA in concert with recruitment by UHRF1 [49–51]. This results in subsequent methylation of the daughter strand based on the pattern of the parent strand, thus maintaining a symmetrical pattern through cell division. However, DMNT1 has reduced binding to 5hmC compared with 5mC [52,53]. Therefore, the conversion of 5mC to 5hmC prevents DMNT1 binding to the parent strand with resulting loss of methylation on the daughter strand. This has been supported by an *in vivo* study revealing passive 5hmC loss in the paternal mouse zygote during preimplantation [54]. Nonetheless, although the model of passive demethylation is supported by a number of groups, it has not been directly shown that the presence of 5hmC in early zygotes equates to a loss of 5mC. The exact mechanism is not as clear, as the UHRF1 cofactor has been shown to bind with equal efficiency to 5mC and 5hmC, therefore raising questions over the DMNT1 maintenance mechanism and the role of 5hmC in preventing such a mechanism from working [55]. Furthermore, a recent study confirmed that while exogenous UHRF1 in embryonic stem (ES) cells improves localization of DMNT1, it does not affect methylation dynamics [56]. There have also been a number of potential active pathways proposed for the DNA demethylation process (Figure 2).

It has been shown that the TET enzymes facilitate the progressive oxidation of 5mC to 5hmC, and subsequently to the recently identified 5fC and 5caC derivatives [57–59]. The decarboxylation of 5caC has been proposed as a candidate route of demethylation, akin to the process of demethylation of thymine to uracil

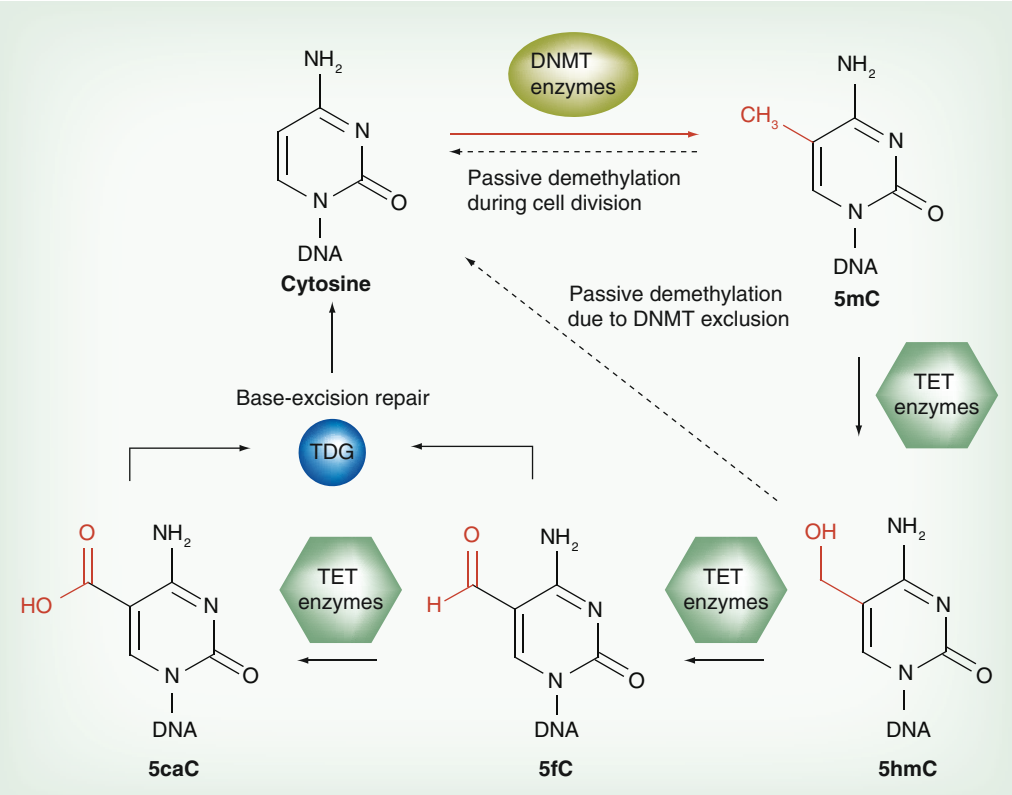

Figure 2. Proposed models of TET-mediated DNA demethylation pathways. Unmodified cytosine bases in a CpG dinucleotides can be directly methylated through the actions of the DNA methyltransferase enzymes. Demethylation is thought to be both passive (during cell division; dashed arrow) as well as active (through a series of enzymatic reactions). TET proteins can oxidize 5mC-modified bases into 5hmC and then onto 5fC and 5caC derivatives. The oxidized 5fC and 5caC bases are thought to provide suitable substrates for rapid demethylation to nonmodified cytosine via TDG-coupled base-excision repair. In contrast with these rapidly turned-over intermediates, 5hmC appears to remain stable in certain parts of the genome.
Black arrows: active methylation; dashed arrows: passive demethylation; red arrow: methylation.
5caC: 5-carboxylcytosine; 5fC: 5-formylcytosine; 5hmC: 5-hydroxymethylcytosine; 5mC: 5-methylcytosine.

in yeast [60], but no such carboxylase has yet been identified in mammals. Alternatively, it has been suggested that demethylation occurs through deamination of 5hmC to 5-hydroxymethyluracil with subsequent base-excision repair, resulting in a nonmodified CpG dinucleotide [61–63]. Once more, the evidence supporting such a pathway has also been called into question. Despite the finding that the AID and APOBEC enzymes can deaminate 5mC *in vitro*, and that they are significantly expressed in murine tissue and primordial germ cells in which demethylation occurs [64,65], they were found to have greater affinity for cytosine binding than for 5mC or 5hmC. Additionally, overexpression of AID/APOBEC did not result in increased detection of deamination products or reduction in 5fC or 5caC [66]. Instead, the most likely mode of active demethylation appears to be regulated by TDG-mediated base-excision repair [57,67–71]. In two recent papers, both 5caC and 5fC were mapped across the genome of mouse ES cells for the first time, with extremely low levels of enrichment observed [70,71]. This is likely due to the fact that these modifications are not stable, but are instead rapidly turned over transient marks of active DNA demethylation. However, upon the reduction of the TDG enzyme by both shRNAi, as well as in a knockout model, ectopic regions of 5fC and 5caC become apparent and were seen over genic and promoter-proximal regions, with a particular enrichment over poised (H3K4me1 but not H3K27ac-marked) enhancer elements. Taken together with the finding that depletion of the TDG enzyme in mouse ES cells leads to accumulation of 5caC to detectable levels, this suggests that this is a likely mechanism for active demethylation events [57]. Nevertheless, it is of note that the detection of 5fC and 5caC in the

paternal pronucleus during mouse preimplantation and the dilution of these factors through replication, similar to that described for 5hmC, may also suggest that these factors could be involved in a passive demethylation process [54,72]. While the involvement of 5hmC in the demethylation process appears highly plausible, the molecular mechanisms through which this occurs have yet to be unequivocally identified. It is clear that 5hmC is not solely an intermediate in this process (unlike the transient marks of 5fC and 5caC) and is likely to have additional unique biological functions.

One such biological role may be associated with marking gene transcription. The conversion of 5mC to 5hmC affects the binding affinity of a number of proteins to DNA. There have been reports of reduced affinity of MBD1, MBD2, MBD4 and MeCP2 to 5hmC [73,74]. These are contrasted with reports of significant affinity by MBD3 and MeCP2 to 5hmC-containing substrates [75,76]. 5hmC binding by MeCP2 has been reported in extracts derived from ES cells, but not in neuronal tissue, while MeCP2 was found to be a 5mC reader in mouse ES cells, neuronal progenitor cells and adult mouse brain tissue [77]. In this analysis, the protein–DNA interactions were dynamic with, for example, neuronal progenitor cell-specific binding of Uhrf2 to 5hmC. Wdr76 preferentially binds to 5hmC and interacts with HELLS, a DNA helicase that has been previously implicated in regulating DNA methylation levels in cells [78,79]. While the exact consequences of altered protein binding have not been elucidated, it may suggest that 5hmC is an important contributor to the control of gene transcription through its ability to differentially interact with nuclear factors, in addition to DNA methylation reprogramming.

5hmC occurs in CpG dinucleotides with an asymmetrical strand bias, with a G-rich preference in human ES cells [43]. In human somatic neural tissue, 5hmC is enriched in promoters and intragenic regions, but is largely absent from intergenic regions [80]. Moreover, 5hmC within gene bodies appears to be preferentially located at exons [34,37,81] and is positively correlated with gene-expression levels [80], which mirrors findings from work on murine neuronal tissue [39]. In addition, further work on both human and mouse ES cells has shown genic 5hmC distribution to be related to gene expression with high 5hmC levels in the gene bodies of high- and intermediate-expressed genes compared with low-expressed genes [39,82]. This may suggest that there is a mechanism that allows TET enzymes to target genes [41], presumably for gene activation, although the causative effect on gene regulation is yet to be shown. TET1, 2 and 3 have now been identified as interaction partners for OGT. Chromatin immunoprecipitation-sequencing experiments demonstrate that OGT binding sites on chromatin are co-occupied by TET enzymes in mouse ES and somatic cells [83–86]. In addition, TET proteins and OGT activity can promote binding of the SET1/COMPASS H3K4 methyltransferase, SETD1A, to chromatin [84]. TET2 knockout in mouse bone marrow leads to decreases in global *N*-acetyl glycosylation and H3K4me3, notably at several key regulators of hematopoiesis. Together, these results suggest a novel pathway by which the TET enzymes themselves may be required for transcriptional activation in the absence of their catalytic function. **Figure 3** illustrates the proposed TET, OGT and SET1 interactions for gene transcription, with typical 5mC and 5hmC profiles associated with the active gene.

Interestingly, in human ES cells, 5hmC has been found to be most abundant at regions of low CpG density, and promoters with high CpG content (e.g., CpG islands) have almost complete absence of 5hmC [43,82]. Study of human ES cells also revealed that 5hmC is located at regions with a GC skew, in which Gs are enriched over Cs in the 5′ end of the region with the converse in the 3′ end [87]. It has previously been proposed that GC skew may occur at sites of replication termination and recombination hotspots [88–90] and, as such, 5hmC may also mark these sites [87].

A major modifier of both global and locus-specific 5hmC profile is tissue type [41]. Interestingly, in both mouse and human tissues, the brain, which has a low proliferative rate, had the highest level of 5hmC, while blood, which undergoes significant proliferation during hematopoiesis, had the lowest [41]. This same study also confirmed loss of 5hmC from cells that have adapted to culture. Together, these findings may reinforce the role of 5hmC in the passive demethylation process, in which the proliferative rate of tissue and cells affects 5hmC levels. This hypothesis is strengthened by the findings of Jin *et al.* who demonstrated with immunohistochemistry the mutual exclusivity of Ki67, a marker of proliferation and 5hmC in tissue [91]. However, the finding of reduced TET expression upon cell culture may suggest that this is not solely related to proliferation [41].

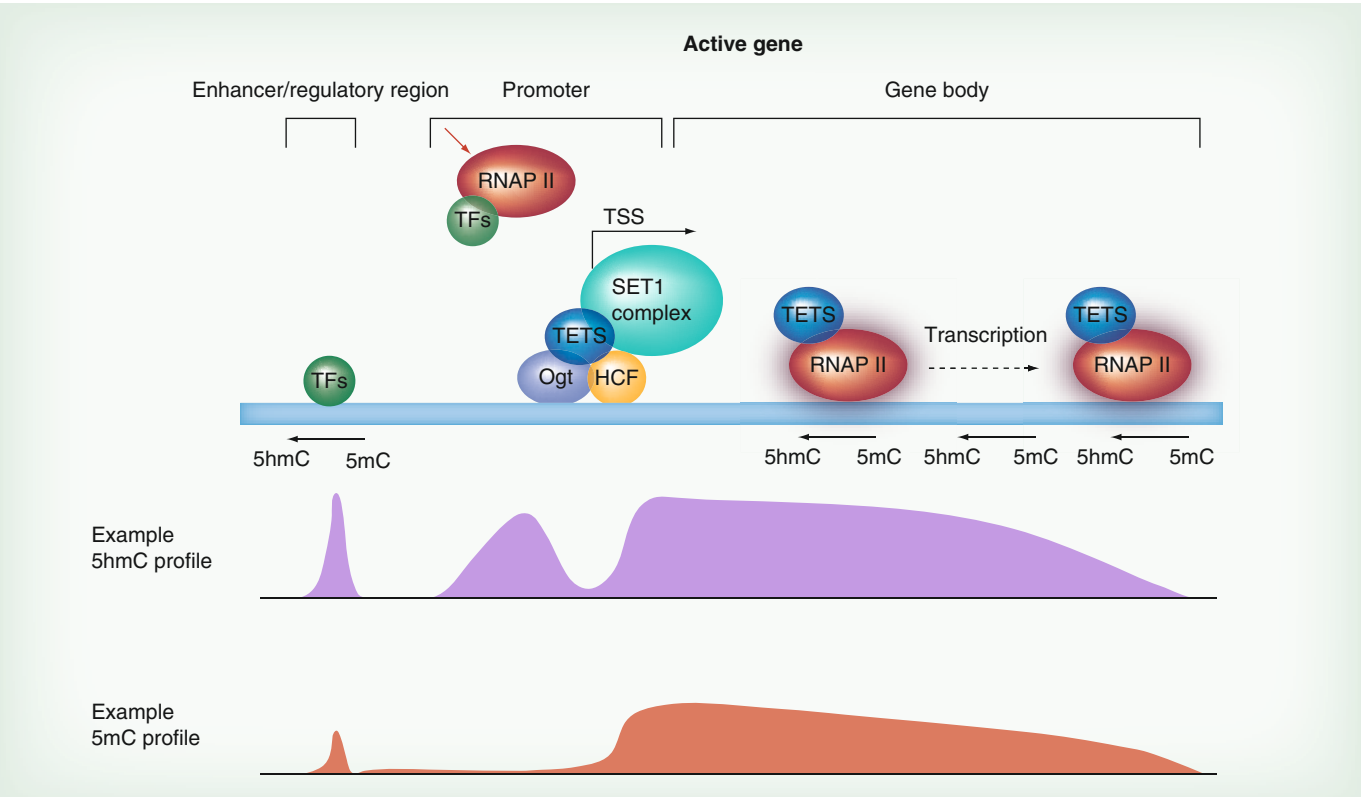

Figure 3. Proposed model of 5-methylcytosine, 5-hydroxymethylcytosine and TET distribution associated with gene transcription. Activation of gene expression can occur through the active demethylation of regions around the TSS (loss of 5mC through a 5hmC intermediate) with elevated promoter and gene body 5hmC levels allowing the binding and elongation of the RNAP-II complex, probably in concert with histone modification changes (not shown). Conversion of 5mC into 5hmC (and further derivatives) is mediated by the TET proteins (blue). TET proteins have been show to interact with OGT, as well as the HCF1, component of the H3K4 methyltransferase SET1/COMPASS complex, resulting in altered chromatin environments (OGT, HCF1 and SET1 complex: yellow). TDG (not shown) is proposed to complete this promoter-specific demethylation through base-excision repair, while regions lacking TDG lead to the accumulation of 5hmC (e.g., gene bodies and enhancers), perhaps by tracking the RNAP II complex in the case of gene bodies. Typical 5hmC (purple) and 5mC (red) profiles at an actively transcribing gene are shown above.
5hmC: 5-hydroxymethylcytosine; 5mC: 5-methylcytosine; TF: Transcription factor; TSS: Transcription start site.

5hmC profiling defines cellular state

Recently, accumulation of 5hmC in the developing mammalian brain has been shown to occur in a partly independent process from methylation at CpGs [92]. The development of adult 5hmC patterns in the brain have also been shown to develop *in utero*, although the 5hmC pattern in the fetal frontal cortex is unique to the stage of fetal development [92]. Furthermore, developmentally downregulated genes show enrichment of 5hmC in the fetal frontal cortex, but not in adults, supporting the correlation of reduced transcription with loss of 5hmC [76,92]. These findings support the role of 5hmC as a marker of cellular state, which has been demonstrated systematically using a nongenotoxic carcinogen (NGC) liver tumor mouse model [93,94].

It is known that the epigenome can be perturbed in response to a group of chemicals called NGCs. Exposure to the NGC phenobarbital (PB) can lead to an increase in the incidence of spontaneously and chemically induced liver tumors in rodents and significantly promotes hepatic tumor incidence in B6C3F1 mice, as well as increasing the size of the tumors themselves [95–99].

PB exposure leads to the expression of genes that depend on activation of the constitutive active androstane receptor. It was recently suggested that PB activates the constitutive active androstane receptor by inhibiting EGF receptor signaling [100]. Recent work in our laboratory shows that the distribution of 5mC/5hmC is highly consistent between untreated individual mice of a similar age; yet subtle changes during liver maturation in a transcriptionally dependent manner occur [93]. Following exposure to PB, there are staged transcriptional responses corresponding to dose exposure that strongly correlate with promoter-proximal region 5hmC levels, suggesting that 5hmC may be

the mechanism that facilitates transcriptional changes [93,94]. Furthermore, reciprocal changes for both 5mC and 5hmC in response to PB suggest that active demethylation may be taking place at each set of these loci, via a 5hmC intermediate. It is possible that these changes in 5mC and 5hmC result from variation in cell populations and cellular heterogeneity in response to stimuli. However, the dynamics of 5hmC turnover appear to be very rapid, implying that these reflect true cell autonomous changes in the epigenome. As such, this work suggests 5hmC profiling can be used as an indicator of cell states during organ maturation and drug-induced responses, and provides novel epigenetic signatures for NGC exposure. Many regions were seen to change dramatically in their 5hmC levels following drug exposure. By focusing on the promoter regions of strongly induced genes, it was observed that these contained elevated levels of 5hmC-marked DNA, as well as increased levels of H3K4me2, a histone modification associated with gene activation [94]. Conversely, these same promoter regions were found to lose 5mC. PB-induced genes gained H3K36me3 in the body of the gene, while losing the repressive histone mark, H3K27me3.

Importantly, this work revealed that reproducible interindividual epigenetic perturbations, namely changes to the hydroxymethylome, can be observed following 24 h of drug dosing and that these changes persist or expand over several months as the length of exposure continues [93]. This raises the exciting possibility that 5hmC is a potential biomarker of exposure to particular environmental stimuli or risk factors in disease development. While 5mC patterns have been demonstrated to be changed in response to environmental stimuli such as smoking [101], 5hmC may offer additional benefit as a marker of specific exposure. This was demonstrated with the PB NGC model, as global 5hmC perturbations alone were sufficient to stratify drug-exposed individuals from the control set, while 5mC and selective histone modifications alone did not always reliably report exposure [93,94]. It would be important to confirm whether these changes remain present and identifiable long after exposure in this model, and whether these observations in a mouse model system are transferable to human disease.

The measurable and reproducible changes in 5hmC in response to a drug could conceivably be used to monitor epigenetic changes in the tissue as an indication of a real-time response. Furthermore, the alterations observed in response to a previously considered 'nonepigenetic modulator' such as PB, could suggest that these profiles may be used to monitor traditional therapies, as well as newer treatments targeted at epigenetic components. One might expect a reversal of the 'cancer signature' to that of the 'normal tissue signature' on successful treatment of a disease in a tissue-specific manner. This would be particularly important in oncological management, where there is a void of sensitive and specific markers of disease response, and patients often undergo repeated ionizing imaging to assess disease burden.

5hmC patterns in cancer

Cancer-related 5hmC pattern changes and possible underlying mechanisms are summarized in **Box 1**. Reduction in 5hmC is well reported in a number of cancers and cell lines corresponding to prostate, breast, colon, lung, brain, liver, kidney and melanoma, compared with the associated normal tissue [41,91,102]. This hypohydroxymethylation may be due to the cellular replication-related loss of 5hmC described above, reflecting the increased proliferative

Box 1. Summary of changes in 5-hydoxymethylcytosine patterns and possible underlying mechanisms seen in cancer.

- Cancer-associated hydroxymethylcytosine patterns
- General loss of 5hmC
- Cancer-specific redistribution with enrichment, particularly at oncogenic gene activators
- Possible mechanisms for 5hmC changes in cancer
- Replication-related passive demethylation and loss of hydroxymethylcytosine
- Misexpression or mutation of TET enzymes, which are responsible for the oxidative conversion of 5mC to 5hmC
- Inhibition of the essential TET cofactor, α-ketoglutarate, through mutation of isocitrate dehydrogenase or other Krebs cycle enzymes, such as fumerate hydratase or succinate dehydrogenase

5mC: 5-methylcytosine; 5hmC: 5-hydroxymethylcytosine.

rate of the cancer compared with normal tissue. However, there was no correlation between levels of 5hmC and 5mC staining in colorectal or prostate tumors [102], suggesting the presence of alternative mechanisms to the general global hypomethylation reported in cancer [103].

Another possible explanation for altered 5hmC patterns in cancer is the misexpression of TET enzymes, which are responsible for 5hmC production. While TET3 has been implicated in the 5mC decrease, and simultaneous 5hmC increase in the male pronucleus upon zygote formation prior to cell division [104–106], there has been no association between TET3 expression and cancer reported so far.

Prior to the rediscovery of 5hmC with the subsequent appreciation of TET involvement in the methylation cycle, TET1 was identified as a fusion partner of the mixed-lineage leukemia in the ten–eleven chromosomal translocation t(10;11)(p12;q23) in rare cases of acute myeloid leukemia (AML) [107,108]. Whereas the role of TET1 in these AML cases is independent of its regulation of 5hmC, subsequent loss of TET1 and changes in 5hmC have been shown in a number of solid tumors. In prostate and breast cancer, TET1 has been shown to inhibit cancer growth and metastases through tissue inhibitors of metalloproteases (TIMPs). Conversely, loss of TET1 is associated with the *de novo* methylation of TIMP2 and TIMP3 and associated gene repression, and correlates with poor survival rates in breast cancer patients [109]. The HMGA2-TET1-HOXA9 pathway is coordinately regulated in breast cancer and has been reported to encompass a prognostic signature for patient survival [110]. Expression of TET1 and HOXA9 suppresses breast tumor growth and metastasis in mouse xenografts. Others have demonstrated that loss of 5hmC was associated with downregulation of TET1 in hepatocellular carcinoma and this was associated with poorer overall survival [111]. Clearly, these studies suggest that the loss of TET1 may have an important role in the development of an aggressive phenotype in cancer.

As well as the alterations of TET1 in cancers, TET2 has been implicated in melanoma progression. Global loss of 5hmC was demonstrated in malignant melanoma and this correlated with loss of TET expression, particularly TET2 [112]. Furthermore, these global changes in 5hmC were associated with Breslow score, mitotic rate, pathological tumor stage and tumor ulceration, all currently used prognostic indicators in malignant melanoma. Loss of 5hmC in melanoma patients was also correlated with poorer survival on Kaplan–Meier survival analysis [112].

A study in pancreatic cancer has shown that there is also a cancer-specific redistribution of 5hmC, with enrichment in particular oncogenic gene activators [113]. The highly tissue-specific and dynamic patterns observed for 5hmC distribution could be vital in defining the tissue of origin of a metastatic cancer, which in turn may aid the overall diagnosis and subsequent treatment regime [8]. These findings, combined with the cell- and tissue-specific distribution of 5hmC, suggest a possible prognostic role for 5hmC in cancer, although multivariate analysis has not been performed to confirm whether 5hmC could be independently significant. Furthermore, overexpression of TET2 in melanoma cells resulted in subsequent partial re-establishment of 5hmC profiles with inhibition of invasion on cell culture-based invasion assays and tumor growth *in vitro* using xenograft models [112]. This suggests possible success with therapeutic manipulation of TET2, but also provides support for 5hmC as a marker of response to treatment.

Despite these changes in TET expression in cancer, TET has not been seen to be frequently mutated in the large-scale sequencing studies of solid malignancies [114–117]. This would suggest that there are mechanisms, other than mutations of the TET genes, which cause these changes. Therefore, further work is required to understand the mechanisms underpinning TET-related 5hmC changes in cancer.

Nonetheless, mutations in TET2 have been demonstrated in all subtypes of myeloid malignancies, and TET2 is the most commonly mutated gene in myelodysplastic syndrome (MDS) [118–123]. These mutations are associated with transformation of MDS to AML, as well as poorer survival in AML patients [118]. Loss-of-function mutations in TET2 have also been demonstrated as a recurrent event in human lymphomas [124]. The breadth of hematological malignancy with *TET2* mutations may support this as an early event in disease development. Moreover, unlike the regulatory control of TET enzymes on 5hmC in murine ES cells to maintain pluripotency [125], it has been shown that this mechanism has the reverse function for control of hematopoietic stem and progenitor cell homeostasis. A number of studies using TET2 mouse models have shown that TET2-null mice have increased hematopoietic stem cell numbers and that these cells are able to self-renew in culture. Furthermore, although these knockout mice are viable, they die early as a result

of hematological malignancy [124,126–128]. These findings, supported by the higher TET2 expression in differentiated blood cells compared with progenitor cells [128], strengthens the argument that the correlation of reduced 5hmC and *TET2* mutation in MDS, and other hematological malignancies, is due to an early event in hematopoiesis, likely to be affecting hematopoietic stem and progenitor cells [124]. Importantly, if these findings are confirmed then it would indicate that disease treatments would need to focus on these hematopoietic stem cells.

The third possible contributing mechanism for the general loss of 5hmC in cancer is through inhibition of TET cofactors. As previously described, TET enzymes are 2OG-dependant, which is produced through catalytic oxidative carboxylation of isocitrate by isocitrate dehydrogenases (IDHs) in the Krebs cycle [129,130]. There are two homologs of this enzyme, IDH1 and IDH2, which catalyze the same reaction. Gain-of-function mutations of both IDH1 and IDH2 have been identified in cancer cells. These mutations produce the R-enantiomer of the oncometabolite 2-hydroxyglutarate (2-HG) [131,132]. Both the R- and S-enantiomers of 2-HG are structurally similar to 2OG and can antagonize the 2OG-dependant reaction [130,133,134], thus inhibiting TET-mediated 5mC to 5hmC conversion. Significant downregulation of IDH2 in melanomas compared with benign nevi has been shown [112]. Furthermore, overexpression of IDH2 in the *BRAF*^{V600E} zebrafish melanoma models significantly increased 5hmC levels, with improved progression-free survival [112], highlights the possible therapeutic benefit of targeting this pathway when mutated.

Recently, through transfection of mutant IDH1 in a leukemia cell line (TF1), IDH has been shown to be oncogenic through the development of growth factor independence and impaired differentiation [135]. Interestingly, knockdown of TET2 recapitulated these findings, suggesting that IDH inhibition of TET may be the mechanism of action. However, paradoxically, while the *IDH1* mutant R-enantiomers, but not the S-enantiomers, resulted in the hallmark leukemic transformations, the S-enantiomers are the greater inhibitor of TET2 [135]. This raises the possibility that other 2OG-dependant pathways, other than inhibition of TETs may be involved. Nonetheless, the changes in growth factor dependence and impaired differentiation induced by *IDH1* mutation were demonstrated to be reversible through pharmacological blockade of 2-HG. While we have indicated that there are potential problems of studying TET and 5hmC in cell culture, and clearly further work is required to understand the subsequent mechanism of tumorigenesis following *IDH* mutation, this does raise the possibility of success for 2-HG-blocking compounds as therapeutic agents.

As well as the association of IDH dysregulation in melanoma, mutations in IDH1 and IDH2 have been seen in hematological, brain, colonic, prostate and thyroid tumors [115,132,136–140]. Interestingly, mutations in IDH1 have been identified in up to 70% of grade 2 and 3 gliomas and secondary glioblastomas, with most of the remainder harboring mutations in IDH2 [138]. These mutations appeared early in tumor development and their presence in both astrocytoma and oligodendroglioma subtypes suggests that these may be early mutations in the neural stem cells, a common theme in these 5hmC/TET/IDH abnormalities. However, while IDH mutations in melanoma and cholangiocarcinoma studies have correlated with loss of 5hmC [112,141], this has not been universally demonstrated in brain tumors. *IDH1* mutation has, however, been linked with a distinct pattern of CpG promoter hypermethylation [142], which could be interpreted as being a result of a failure in the conversion of 5mC to 5hmC through TET inhibition. In addition, some brain tumors have been shown to have prognostically lower 5hmC levels than normal brain tissue. Despite this, there has been no correlation with IDH mutations and 5hmC levels in these cancers to date [91,143–145]. It therefore raises the possibility that IDH mutations may drive oncogenesis through alternative pathways to 5hmC. There are a number of other 2OG-dependant pathways, including the hypoxia-inducible factor propyl-hydroxylase [146] and JMJC domain-containing histone demethylases [147], which could be targeted to promote oncogenesis. Furthermore, mutations in succinate dehydrogenase and fumarate hydratase identified in a number of human cancers including pheochromocytoma, paraganglioma and papillary renal cell carcinoma, result in the accumulation of succinate and fumerate, respectively [148–151]. These also inhibit 2OG-dependant deoxygenases, and it is only recently that the potential impact on TET activity has been appreciated with alterations of genome-wide DNA methylation being identified [152,153]. The complete extent of these changes has not been investigated, neither has the role of 5hmC. These findings emphasize the likely importance of interactions between the Krebs cycle, epigenomic regulation and cancer.

Future perspective & clinical applicability

Four years following the rediscovery of 5hmC, there has been significant reassessment and progress in our knowledge of methylation, hydroxymethylation and associated changes in disease states such as cancer. While the interactions of these are complex, and likely tissue-specific, there is clear potential for clinical translation of some of these findings following more in-depth study. These proposed potential strategies for the use of 5hmC as a biomarker could help us to achieve personalized medicine in which we can stratify high-risk patients, prognosticate risk of recurrence, predict sensitivity to treatment and monitor therapeutic response (Figure 4). These possible applications will only be realized if there is further large-scale study of 5hmC profiling in cancer at diagnosis and sequentially throughout therapeutic intervention to define subtype-specific patterns of 5hmC, thus allowing us to interpret the heterogenous outcomes of patients with the same diseases. These studies would also have to define and evade epigenetic heterogeneity, which we know is a significant problem at a genetic level [154]. Furthermore, while predictive and prognostic markers are easily applicable to tumor samples at the time of extirpative surgery or diagnostic biopsy, the translation of 5hmC as a marker of therapeutic response is more complex. Currently, sequential tumor sampling through therapeutic intervention is not routine. This is for two reasons. First, as there are no reliable markers of response, there is no benefit to subjecting the patient to potentially harmful intervention. Second, repeated biopsy may be poorly tolerated by these patients. As such, the use of tissue 5hmC profiles as a dynamic measure of response will only be adopted if significant benefit to the patient can be demonstrated, non- (or minimally) invasive measurements can be identified or biopsy techniques can be made more tolerable.

Despite this, there is definite potential to add to the growing perception that NGS studies can make significant contributions to personalized medicine. There is certainly mounting evidence that 5hmC is involved in DNA methylation reprogramming, however, the exact mechanism by which this occurs still needs to be confirmed. It would also appear that 5hmC has a discrete role aside from a demethylation intermediate, as the mark persists at select loci at detectable (albeit low) levels, while the downstream 5fC and 5caC intermediates are rapidly turned over by base-excision repair [30]. It would also appear that there are significant correlations between 5hmC profiles and gene expression. However, the identification of the mechanisms in which this happens need to be explored. Irrespective of the mechanism, the tissue- and sample-specific profiles of 5hmC provide a unique opportunity for monitoring cell state, which may be highly applicable to archived material and therapeutic screens. However, although there has been progress in profiling 5mC from low-quantity and highly fragmented DNA as a result of formalin fixation and paraffin-embedded tissue [155], there has yet to be any study of 5hmC profiles in archival tissue or study of the effect of such tissue storage on 5hmC patterns, and this is clearly needed. This is especially important

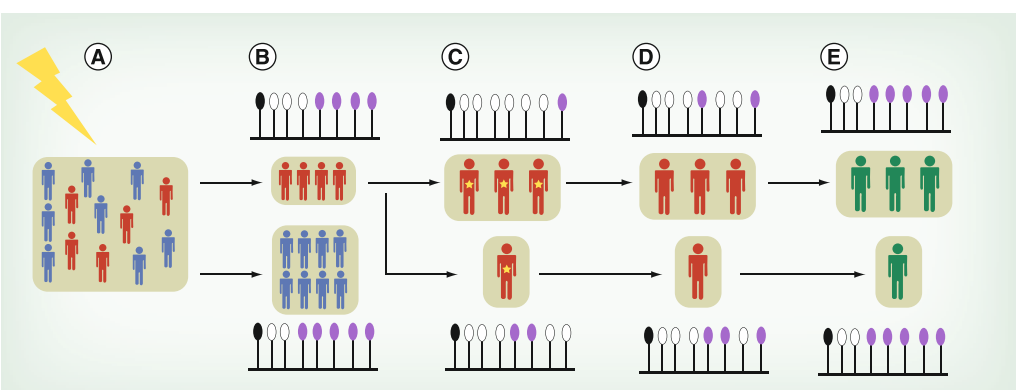

Figure 4. Potential model of 5-hydroxymethylcytosine biomarker use in the diagnosis and management of cancer. (A) Cohort exposed to risk factor for disease. **(B)** Red population have 5-hydroxymethylcytosine (5hmC) signature suggestive of sensitivity to this risk factor (black lollipops: 5-methylcytosine-marked CpGs; purple lollipops: 5hmC marked; white lollipops: unmarked CpGs). These patients can undergo close surveillance to allow early detection or pre-emptive treatment, if available. Blue patients can be reassured of low risk. **(C)** Patients develop cancer and can be stratified into treatment-sensitive or treatment-resistant groups to allow personalized therapy. **(D)** Response to treatment can be followed. **(E)** Return to 'normal' 5hmC tissue signature on cure.

if fixation methods result in general oxidative damage to DNA. If possible, reliable profiling of archival tissue would provide a vast resource of well-annotated clinical samples to aid biomarker discovery.

The roles of TET enzymes and IDH obviously stretch beyond regulation of methylation and hydroxymethylation and teasing out the interactions will be interesting. In parallel to answering these mechanistic questions, such knowledge would allow the development of epigenetic drugs, which could reinstate normal 5hmC patterns, either globally or in a locus-specific manner. The substantial progress thus far gives promise of many exciting developments in this field over the coming years. Ultimately, the analysis of global 5hmC patterns, when combined with other epigenetic modifications and transcriptional information, will lead to a more refined understanding of tumor formation and stratification of specific cancer subtypes.

Acknowledgements

The authors thank D Sproul (MRC Human Genetics Unit, Edinburgh, UK) for invaluable comments on the manuscript. The authors also apologize to colleagues whose work could not be cited due to space limitations.

Financial & competing interests disclosure

A Laird is supported by the Medical Research Council Scottish Clinical Pharmacology and Pathology Programme, The Royal College of Surgeons of Edinburgh Robertson's Trust and The Melville Trust for the Care and Cure of Cancer. J Thomson is supported by the MARCAR project. Work in RR Meehan's laboratory is supported by the Medical Research Council, the BBSRC and by the Innovative Medicine Initiative Joint Undertaking (IMI JU) under grant agreement number 115001 (MARCAR project). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Discovery & technology

- Standard bisulphite conversion and methylation-specific enzymatic digestion for the elucidation of methylation status cannot differentiate between 5-methylcytosine (5mC) and 5-hydroxymethylation (5hmC).
- Antibody-mediated 5hmC-enrichment techniques, glucosylation and purification, or selective restriction enzyme digestion can allow effective differential quantification of 5mC and 5hmC.

Function & distribution

- 5hmC is involved in DNA methylation reprogramming. However, the exact mechanism, whether passive or active, still needs to be confirmed.
- 5hmC is likely to have a distinct role in marking and contributing to control of gene transcription.

Hydroxymethylcytosine profiling defines cellular state

- Global 5hmC levels are tissue-specific. 5hmC profiles are highly conserved in the livers of age-matched mice and are transcription-dependant.
- In the nongenotoxic carcinogen mouse model, global 5hmC perturbations alone were sufficient to stratify drug-exposed individuals from the control set, while 5mC and selective histone modifications did not always report exposure.

Hydroxymethylcytosine patterns in cancer

- General loss and cancer-specific redistribution of 5hmC has been shown. This may be through replication-dependant demethylation, TET inhibition or mutation, or metabolic manipulation by alteration of Krebs cycle enzymes.

Clinical applicability

- Tissue- and cell-specific 5hmC profiles provide a unique opportunity to monitor cellular state and show potential for translation to the clinic as diagnostic, prognostic or predictive biomarkers.
- Understanding the role of 5hmC, TETs and isocitrate by isocitrate dehydrogenase in carcinogenesis may provide novel therapeutic opportunities.

References

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1 Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. *Genes Dev.* 23(7), 781–783 (2009).

2 Bird A. Perceptions of epigenetics. *Nature* 447(7143), 396–398 (2007).

3 Wongtawan T, Taylor JE, Lawson KA, Wilmut I, Pennings S. Histone H4K20me3 and HP1alpha are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J. Cell. Sci.* 124(Pt 11), 1878–1890 (2011).

4 Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447(7143), 425–432 (2007).

5 Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science* 329(5987), 78–82 (2010).

6 Timp W, Feinberg AP. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. *Nat. Rev. Cancer* 13(7), 497–510 (2013).

7 Hansen KD, Timp W, Bravo HC *et al.* Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* 43(8), 768–775 (2011).

8 Sproul D, Meehan RR. Genomic insights into cancer-associated aberrant CpG island hypermethylation. *Brief Funct. Genomics* 12(3), 174–190 (2013).

9 Sproul D, Kitchen RR, Nestor CE *et al.* Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol.* 13(10), R84 (2012).

10 Reddington JP, Perricone SM, Nestor CE *et al.* Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol.* 14(3), R25 (2013) (Epub ahead of print).

11 Cedar H, Bergman Y. Programming of DNA methylation patterns. *Annu. Rev. Biochem.* 81, 97–117 (2012).

12 Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99(3), 247–257 (1999).

13 Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69(6), 915–926 (1992).

14 Ashe A, Morgan DK, Whitelaw NC *et al.* A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol.* 9(12), R182 (2008).

15 Reddington JP, Pennings S, Meehan RR. Non-canonical functions of the DNA methylome in gene regulation. *Biochem. J.* 451(1), 13–23 (2013).

16 Dunican DS, Ruzov A, Hackett JA, Meehan RR. xDnmt1 regulates transcriptional silencing in pre-MBT *Xenopus* embryos independently of its catalytic function. *Development* 135(7), 1295–1302 (2008).

17 Hervouet E, Vallette FM, Cartron PF. Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. *Epigenetics* 4(7), 487–499 (2009).

18 Riggs AD. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* 14(1), 9–25 (1975).

19 Bestor T, Laudano A, Mattaliano R, Ingram V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J. Mol. Biol.* 203(4), 971–983 (1988).

20 Ooi SK, Bestor TH. The colorful history of active DNA demethylation. *Cell* 133(7), 1145–1148 (2008).

21 Tahiliani M, Koh KP, Shen Y *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324(5929), 930–935 (2009).

▪▪ **Identification of TET enzymes that convert 5-methylcytosine to 5-hydroxymethylcytosine.**

22 Koh KP, Rao A. DNA methylation and methylcytosine oxidation in cell fate decisions. *Curr. Opin. Cell. Biol.* 25(2), 152–161 (2013).

23 Shen L, Zhang Y. 5-Hydroxymethylcytosine: generation, fate, and genomic distribution. *Curr. Opin. Cell. Biol.* 25(3), 289–296 (2013).

24 Wyatt GR, Cohen SS. A new pyrimidine base from bacteriophage nucleic acids. *Nature* 170(4338), 1072–1073 (1952).

25 Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem. J.* 126(4), 781–790 (1972).

26 Kothari RM, Shankar V. 5-methylcytosine content in the vertebrate deoxyribonucleic acids: species specificity. *J. Mol. Evol.* 7(4), 325–329 (1976).

27 Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324(5929), 929–930 (2009).

▪▪ **Identification of mammalian 5-hydroxymethylcytosine.**

28 Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS ONE* 5(1), e8888 (2010).

29 Nestor C, Ruzov A, Meehan R, Dunican D. Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *Biotechniques* 48(4), 317–319 (2010).

30 Munzel M, Globisch D, Carell T. 5-hydroxymethylcytosine, the sixth base of the genome. *Angew. Chem. Int. Ed. Engl.* 50(29), 6460–6468 (2011).

31 Clark SJ, Statham A, Stirzaker C, Molloy PL, Frommer M. DNA methylation: bisulphite modification and analysis. *Nat. Protoc.* 1(5), 2353–2364 (2006).

32 Clark SJ, Harrison J, Paul CL, Frommer M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* 22(15), 2990–2997 (1994).

33 Eads CA, Danenberg KD, Kawakami K *et al.* MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res.* 28(8), e32 (2000).

34 Pastor WA, Pape UJ, Huang Y *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473(7347), 394–397 (2011).

35 Weber M, Davies JJ, Wittig D *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* 37(8), 853–862 (2005).

36 Weber M, Hellmann I, Stadler MB *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39(4), 457–466 (2007).

37 Ficz G, Branco MR, Seisenberger S *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473(7347), 398–402 (2011).

38 Huang Y, Pastor WA, Zepeda-Martinez JA, Rao A. The anti-CMS technique for genome-wide mapping of 5-hydroxymethylcytosine. *Nat. Protoc.* 7(10), 1897–1908 (2012).

39 Song CX, Szulwach KE, Fu Y *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* 29(1), 68–72 (2011).

40 Kinney SM, Chin HG, Vaisvila R *et al.* Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. *J. Biol. Chem.* 286(28), 24685–24693 (2011).

41 Nestor CE, Ottaviano R, Reddington J *et al.* Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.* 22(3), 467–477 (2012).

▪ **Potential of 5-hydroxymethylcytosine profile as a tissue identifier.**

42 Booth MJ, Branco MR, Ficz G *et al.* Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* 336(6083), 934–937 (2012).

43 Yu M, Hon GC, Szulwach KE *et al.* Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* 149(6), 1368–1380 (2012).

44 Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. *Annu. Rev. Microbiol.* 62, 235–251 (2008).

45 Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466(7310), 1129–1133 (2010).

46 Freudenberg JM, Ghosh S, Lackford BL *et al.* Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. *Nucleic Acids Res.* 40(8), 3364–3377 (2012).

47 Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.* 25(23), 2436–2452 (2011).

48 Guo JU, Su Y, Zhong C, Ming GL, Song H. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle* 10(16), 2662–2668 (2011).

49 Zhang J, Gao Q, Li P *et al.* S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. *Cell Res.* 21(12), 1723–1739 (2011).

50 Sharif J, Muto M, Takebayashi S *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450(7171), 908–912 (2007).

51 Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X. The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* 455(7214), 826–829 (2008).

52 Valinluck V, Sowers LC. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res.* 67(3), 946–950 (2007).

53 Hashimoto H, Liu Y, Upadhyay AK *et al.* Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* 40(11), 4841–4849 (2012).

54 Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* 334(6053), 194 (2011).

55 Frauer C, Hoffmann T, Bultmann S *et al.* Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS ONE* 6(6), e21306 (2011).

56 Oda M, Oxley D, Dean W, Reik W. Regulation of lineage specific DNA hypomethylation in mouse trophectoderm. *PLoS ONE* 8(6), e68846 (2013).

57 He YF, Li BZ, Li Z *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333(6047), 1303–1307 (2011).

58 Ito S, Shen L, Dai Q *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333(6047), 1300–1303 (2011).

59 Schiesser S, Hackner B, Pfaffeneder T *et al.* Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew. Chem. Int. Ed. Engl.* 51(26), 6516–6520 (2012).

60 Smiley JA, Kundracik M, Landfried DA, Barnes VR Sr, Axhemi AA. Genes of the thymidine salvage pathway: thymine-7-hydroxylase from a *Rhodotorula glutinis* cDNA library and iso-orotate decarboxylase from *Neurospora crassa*. *Biochim. Biophys. Acta* 1723(1–3), 256–264 (2005).

61 Rusmintratip V, Sowers LC. An unexpectedly high excision capacity for mispaired 5-hydroxymethyluracil in human cell extracts. *Proc. Natl Acad. Sci. USA* 97(26), 14183–14187 (2000).

62 Cortellino S, Xu J, Sannai M *et al.* Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146(1), 67–79 (2011).

63 Hashimoto H, Hong S, Bhagwat AS, Zhang X, Cheng X. Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res.* 40(20), 10203–10214 (2012).

64 Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* 279(50), 52353–52360 (2004).

65 Popp C, Dean W, Feng S *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* 463(7284), 1101–1105 (2010).

66 Nabel CS, Jia H, Ye Y *et al.* AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.* 8(9), 751–758 (2012).

67 Gong Z, Zhu JK. Active DNA demethylation by oxidation and repair. *Cell Res.* 21(12), 1649–1651 (2011).

68 Raiber EA, Beraldi D, Ficz G *et al.* Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol.* 13(8), R69 (2012).

69 Zhang L, Lu X, Lu J *et al.* Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat. Chem. Biol.* 8(4), 328–330 (2012).

70 Shen L, Wu H, Diep D *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* 153(3), 692–706 (2013).

71 Song CX, Szulwach KE, Dai Q *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* 153(3), 678–691 (2013).

72 Inoue A, Shen L, Dai Q, He C, Zhang Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell Res.* 21(12), 1670–1676 (2011).

73 Jin SG, Kadam S, Pfeifer GP. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* 38(11), e125 (2010).

74 Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res.* 32(14), 4100–4108 (2004).

75 Yildirim O, Li R, Hung JH *et al.* Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* 147(7), 1498–1510 (2011).

76 Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* 151(7), 1417–1430 (2012).

77 Spruijt CG, Gnerlich F, Smits AH *et al.* Dynamic readers for 5-(hydroxy) methylcytosine and its oxidized derivatives. *Cell* 152(5), 1146–1159 (2013).

78 Meehan RR, Pennings S, Stancheva I. Lashings of DNA methylation, forkfuls of chromatin remodeling. *Genes Dev.* 15(24), 3231–3236 (2001).

79 Dennis K, Fan T, Geiman T, Yan Q, Muegge K. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.* 15(22), 2940–2944 (2001).

80 Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res.* 39(12), 5015–5024 (2011).

81 Williams K, Christensen J, Pedersen MT *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 473(7347), 343–348 (2011).

82 Szulwach KE, Li X, Li Y *et al.* Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet.* 7(6), e1002154 (2011).

83 Shi FT, Kim H, Lu W *et al.* Ten-eleven translocation 1 (tet1) is regulated by O-linked *N*-acetylglucosamine transferase (OGT) for target gene repression in mouse embryonic stem cells. *J. Biol. Chem.* 288(29), 20776–20784 (2013).

84 Deplus R, Delatte B, Schwinn MK *et al.* TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* 32(5), 645–655 (2013).

85 Vella P, Scelfo A, Jammula S *et al.* Tet proteins connect the O-linked N-acetylglucosamine transferase OGT to chromatin in embryonic stem cells. *Mol. Cell* 49(4), 645–656 (2013).

86 Chen Q, Chen Y, Bian C, Fujiki R, Yu X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* 493(7433), 561–564 (2013).

87 Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* 12(6), R54 (2011).

88 Brodie Of Brodie EB, Nicolay S, Touchon M *et al.* From DNA sequence analysis to modeling replication in the human genome. *Phys. Rev. Lett.* 94(24), 248103 (2005).

89 Touchon M, Nicolay S, Audit B *et al.* Replication-associated strand asymmetries in mammalian genomes: toward detection of replication origins. *Proc. Natl Acad. Sci. USA* 102(28), 9836–9841 (2005).

90 Huvet M, Nicolay S, Touchon M *et al.* Human gene organization driven by the coordination of replication and transcription. *Genome Res.* 17(9), 1278–1285 (2007).

91 Jin SG, Jiang Y, Qiu R *et al.* 5-hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* 71(24), 7360–7365 (2011).

▪ **Defines loss of 5-hydroxymethylation in multiple cancers.**

92 Lister R, Mukamel EA, Nery JR *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* 341(6146), 1237905 (2013).

93 Thomson JP, Hunter JM, Lempiainen H *et al.* Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. *Nucleic Acids Res.* 41(11), 5639–5654 (2013).

▪▪ **Defines 5-hydroxymethylation as a dynamic marker of exposure to the nongenotoxic carcinogen phenobarbital and liver development, and highlights the potential use of 5-hydroxymethylation as a marker of exposure to environmental stimuli or therapeutic intervention.**

94 Thomson JP, Lempiainen H, Hackett JA *et al.* Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol.* 13(10), R93 (2012).

95 Peraino C, Fry RJ, Staffeldt E. Reduction and enhancement by phenobarbital of hepatocarcinogenesis induced in the rat by 2-acetylaminofluorene. *Cancer Res.* 31(10), 1506–1512 (1971).

96 Becker FF. Morphological classification of mouse liver tumors based on biological characteristics. *Cancer Res.* 42(10), 3918–3923 (1982).

97 Bachman AN, Phillips JM, Goodman JI. Phenobarbital induces progressive patterns of GC-rich and gene-specific altered DNA methylation in the liver of tumor-prone B6C3F1 mice. *Toxicol. Sci.* 91(2), 393–405 (2006).

98 Aydinlik H, Nguyen TD, Moennikes O, Buchmann A, Schwarz M. Selective pressure during tumor promotion by phenobarbital leads to clonal outgrowth of beta-catenin-mutated mouse liver tumors. *Oncogene* 20(53), 7812–7816 (2001).

99 Lee GH. Paradoxical effects of phenobarbital on mouse hepatocarcinogenesis. *Toxicol. Pathol.* 28(2), 215–225 (2000).

100 Mutoh S, Sobhany M, Moore R *et al.* Phenobarbital indirectly activates the constitutive active androstane receptor (CAR) by inhibition of epidermal growth factor receptor signaling. *Sci. Signal.* 6(274), ra31 (2013).

101 Lee KW, Pausova Z. Cigarette smoking and DNA methylation. *Front. Genet.* 4, 132 (2013).

102 Haffner MC, Chaux A, Meeker AK *et al.* Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget* 2(8), 627–637 (2011).

103 Ehrlich M. DNA hypomethylation in cancer cells. *Epigenomics* 1(2), 239–259 (2009).

104 Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. *Cell* 146(6), 866–872 (2011).

105 Iqbal K, Jin SG, Pfeifer GP, Szabo PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl Acad. Sci. USA* 108(9), 3642–3647 (2011).

106 Wossidlo M, Nakamura T, Lepikhov K *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* 2, 241 (2011).

107 Ono R, Taki T, Taketani T, Taniwaki M, Kobayashi H, Hayashi Y. LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). *Cancer Res.* 62(14), 4075–4080 (2002).

108 Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* 17(3), 637–641 (2003).

109 Hsu CH, Peng KL, Kang ML *et al.* TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. *Cell Rep.* 2(3), 568–579 (2012).

110 Sun M, Song CX, Huang H *et al.* HMGA2/TET1/HOXA9 signaling pathway regulates breast cancer growth and metastasis. *Proc. Natl Acad. Sci. USA* 110(24), 9920–9925 (2013).

▪ **Highlights the utility of 5-hydroxymethylation as a diagnostic and prognostic marker in cancer, associating this mechanistically with TET1 expression.**

111 Liu C, Liu L, Chen X *et al.* Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1. *PLoS ONE* 8(5), e62828 (2013).

112 Lian CG, Xu Y, Ceol C *et al.* Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* 150(6), 1135–1146 (2012).

▪▪ **Evidence of 5-hydroxymethylation as a diagnostic and prognostic marker in melanoma, implicating TET and isocitrate dehydrogenase mechanistically. Also highlights the potential therapeutic opportunities for 5-hydroxymethylation and TET enzymes.**

113 Bhattacharyya S, Yu Y, Suzuki M *et al.* Genome-wide hydroxymethylation tested using the HELP-GT assay shows redistribution in cancer. *Nucleic Acids Res.* 41(16), e157 (2013).

114 Wood LD, Parsons DW, Jones S *et al.* The genomic landscapes of human breast and colorectal cancers. *Science* 318(5853), 1108–1113 (2007).

115 Sjoblom T, Jones S, Wood LD *et al.* The consensus coding sequences of human breast and colorectal cancers. *Science* 314(5797), 268–274 (2006).

116 Kan Z, Jaiswal BS, Stinson J *et al.* Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* 466(7308), 869–873 (2010).

117 Berger MF, Lawrence MS, Demichelis F *et al.* The genomic complexity of primary human prostate cancer. *Nature* 470(7333), 214–220 (2011).

118 Abdel-Wahab O, Mullally A, Hedvat C *et al.* Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* 114(1), 144–147 (2009).

119 Delhommeau F, Dupont S, Della Valle V *et al.* Mutation in TET2 in myeloid cancers. *N. Engl. J. Med.* 360(22), 2289–2301 (2009).

120 Jankowska AM, Szpurka H, Tiu RV *et al.* Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood* 113(25), 6403–6410 (2009).

121 Kohlmann A, Grossmann V, Klein HU *et al.* Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in TET2, CBL, RAS, and RUNX1. *J. Clin. Oncol.* 28(24), 3858–3865 (2010).

122 Langemeijer SM, Kuiper RP, Berends M *et al.* Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat. Genet.* 41(7), 838–842 (2009).

123 Tefferi A, Lim KH, Abdel-Wahab O *et al.* Detection of mutant TET2 in myeloid malignancies other than myeloproliferative neoplasms: CMML, MDS, MDS/MPN and AML. *Leukemia* 23(7), 1343–1345 (2009).

124 Quivoron C, Couronne L, Della Valle V *et al.* TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* 20(1), 25–38 (2011).

125 Koh KP, Yabuuchi A, Rao S *et al.* Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 8(2), 200–213 (2011).

126 Ko M, Bandukwala HS, An J *et al.* Ten-eleven-translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl Acad. Sci. USA* 108(35), 14566–14571 (2011).

127 Li Z, Cai X, Cai CL *et al.* Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* 118(17), 4509–4518 (2011).

128 Moran-Crusio K, Reavie L, Shih A *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* 20(1), 11–24 (2011).

129 Reitman ZJ, Jin G, Karoly ED *et al.* Profiling the effects of isocitrate dehydrogenase 1 and 2 mutations on the cellular metabolome. *Proc. Natl Acad. Sci. USA* 108(8), 3270–3275 (2011).

130 Xu W, Yang H, Liu Y *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* 19(1), 17–30 (2011).

131 Dang L, White DW, Gross S *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* 462(7274), 739–744 (2009).

132 Ward PS, Patel J, Wise DR *et al.* The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* 17(3), 225–234 (2010).

133 Chowdhury R, Yeoh KK, Tian YM *et al.* The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep.* 12(5), 463–469 (2011).

134 Koivunen P, Lee S, Duncan CG *et al.* Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. *Nature* 483(7390), 484–488 (2012).

135 Losman JA, Looper RE, Koivunen P *et al.* (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science* 339(6127), 1621–1625 (2013).

136 Abdel-Wahab O, Manshouri T, Patel J *et al.* Genetic analysis of transforming events that convert chronic myeloproliferative neoplasms to leukemias. *Cancer Res.* 70(2), 447–452 (2010).

137 Parsons DW, Jones S, Zhang X *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* 321(5897), 1807–1812 (2008).

138 Yan H, Parsons DW, Jin G *et al.* IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* 360(8), 765–773 (2009).

139 Kang MR, Kim MS, Oh JE *et al.* Mutational analysis of IDH1 codon 132 in glioblastomas and other common cancers. *Int. J. Cancer* 125(2), 353–355 (2009).

140 Murugan AK, Bojdani E, Xing M. Identification and functional characterization of isocitrate dehydrogenase 1 (IDH1) mutations in thyroid cancer. *Biochem. Biophys. Res. Commun.* 393(3), 555–559 (2010).

141 Wang P, Dong Q, Zhang C *et al.* Mutations in isocitrate dehydrogenase 1 and 2 occur frequently in intrahepatic cholangiocarcinomas and share hypermethylation targets with glioblastomas. *Oncogene* 32(25), 3091–3100 (2013).

142 Noushmehr H, Weisenberger DJ, Diefes K *et al.* Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* 17(5), 510–522 (2010).

143 Kraus TF, Globisch D, Wagner M *et al.* Low values of 5-hydroxymethylcytosine (5hmC), the "sixth base," are associated with anaplasia in human brain tumors. *Int. J. Cancer* 131(7), 1577–1590 (2012).

144 Muller T, Gessi M, Waha A *et al.* Nuclear exclusion of TET1 is associated with loss of 5-hydroxymethylcytosine in IDH1 wild-type gliomas. *Am. J. Pathol.* 181(2), 675–683 (2012).

145 Orr BA, Haffner MC, Nelson WG, Yegnasubramanian S, Eberhart CG. Decreased 5-hydroxymethylcytosine is associated with neural progenitor phenotype in normal brain and shorter survival in malignant glioma. *PLoS ONE* 7(7), e41036 (2012).

146 Bruick RK, McKnight SL. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* 294(5545), 1337–1340 (2001).

147 Tsukada Y, Fang J, Erdjument-Bromage H *et al.* Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439(7078), 811–816 (2006).

148 Baysal BE, Ferrell RE, Willett-Brozick JE *et al.* Mutations in SDHD, a mitochondrial complex II gene, in hereditary paraganglioma. *Science* 287(5454), 848–851 (2000).

149 Hao HX, Khalimonchuk O, Schraders M *et al.* *SDH5*, a gene required for flavination of succinate dehydrogenase, is mutated in paraganglioma. *Science* 325(5944), 1139–1142 (2009).

150 Bayley JP, Kunst HP, Cascon A *et al.* *SDHAF2* mutations in familial and sporadic paraganglioma and phaeochromocytoma. *Lancet Oncol.* 11(4), 366–372 (2010).

151 Tomlinson IP, Alam NA, Rowan AJ *et al.* Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat. Genet.* 30(4), 406–410 (2002).

152 Letouze E, Martinelli C, Loriot C *et al.* SDH mutations establish a hypermethylator phenotype in paraganglioma. *Cancer Cell* 23(6), 739–752 (2013).

153 Xiao M, Yang H, Xu W *et al.* Inhibition of alpha-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. *Genes Dev.* 26(12), 1326–1338 (2012).

154 Gerlinger M, Rowan AJ, Horswell S *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366(10), 883–892 (2012).

155 Thirlwell C, Eymard M, Feber A *et al.* Genome-wide DNA methylation analysis of archival formalin-fixed paraffin-embedded tissue using the Illumina Infinium HumanMethylation27 BeadChip. *Methods* 52(3), 248–254 (2010).